

0923

#6



OIPE

RAW SEQUENCE LISTING

DATE: 09/23/2002

PATENT APPLICATION: US/10/005,337A

TIME: 10:56:26

Input Set : A:\Substitute Sequence Listing.txt

Output Set: N:\CRF3\09232002\J005337A.raw

ENTERED

```

3 <110> APPLICANT: BENOIT, Patrick
4   SCHWARTZ, Bertrand
5   BRANELLEC, Didier
6   CHIEN, Kenneth R.
8 <120> TITLE OF INVENTION: SEQUENCES UPSTREAM OF THE CARP GENE, VECTORS CONTAINING
9   THEM AND USES THEREOF
11 <130> FILE REFERENCE: 03806.0530-00000
13 <140> CURRENT APPLICATION NUMBER: US 10/005,337A
15 <141> CURRENT FILING DATE: 2001-12-07
17 <150> PRIOR APPLICATION NUMBER: US 60/251,582
18 <151> PRIOR FILING DATE: 2000-12-07
20 <160> NUMBER OF SEQ ID NOS: 5
22 <170> SOFTWARE: PatentIn Ver. 2.1
24 <210> SEQ ID NO: 1
25 <211> LENGTH: 2358
26 <212> TYPE: DNA
27 <213> ORGANISM: Mus musculus
29 <400> SEQUENCE: 1
30 ggatcctttc atgtttaaca atatcaaccc taaccaaggg ggaacagcct gcctgacagt 60
31 ggctttgcca cccatgaata cttoctagtc tagtccgttt gtgaaactca gcccatccca 120
32 acacttctgc aagccccatc ctctacaagg tgctcattgg gaatttctctg gagcttctct 180
33 ttcaggatca gcctgattct agggcagcag ttctcaacct gggggcctcg accccttttg 240
34 gggaatcaaa cgacccttta caggggtcac atatcatcta tctatatgt caggatatta 300
35 cattacgatt cgtaacagta gcaaaattac aggtatgaaa tagcaatgaa ataattttat 360
36 gattgaaggt caccacaaca tgaggccgcc acactgttct agagaaaaat cacctgggtg 420
37 gggaaagggt tgggaaagcc tttctgtcca ttcttcattc ttcaaagtga tgtgttcaca 480
38 gaaagccttt cagctgttct gctggggctc ttagtaagtc tgagtaggaa ctgtatgtac 540
39 cagggtctgt tcttatgggt ggagccaaga cgcctcgtgg gtggagcgaa gacgcaacct 600
40 caccctctag ctctgcatoc atagcaagta gcctaattgt tctgtgtcta ggtgtcatct 660
41 ctgtgaatcg agatccttgg ccttgcttga attagggagg cacaaaatac tcagagattc 720
42 aagactgctc agcagcccg agtcccttcc caaaggaaag gtctcaactc tcagcccccc 780
43 ttagctctga gtcaggcctg gaacaaacgg ccacaggaat gagaaaagct gccatagctg 840
44 cttgtcactt caagagggtca aagaaaatag tgtaaccat gaaaacgaga agaccaacag 900
45 ttatccattg atagcgtctc aggacagata ggacagagag aacactagga gagggggaacc 960
46 cagcaaggac aaggtattag tgtgttggtt ttcagggcaa tgtcttgtag tgaagattct 1020
47 agaaacacaa tttgctgggt gaacagctga agtgggggtg ggggttcttac cccatgttca 1080
48 tgggaagggt agtgaggaga gacagatata tgatggccag cataacaaac atacacaaca 1140
49 ccctaattaa cacttccctc ttctactgac accccttca ctctcctctt tcataaaaaa 1200
50 taaaaaaagt attttatgtg gctcttacga tagaatcttt cctcgaacta taaaaagatc 1260
51 taaatattta tatttttcac attttaatat cttagcgatg acaagccaga aacaagtatt 1320
52 ttttgccctc ctcaacagca aagcttgggg cctttttgtt tccgtgttag gaatagaaca 1380
53 cgagagcccc gtgtatctag gcagatgctc tatcattagc ccatgagtct ccagcctcag 1440
54 acgcacattt ttctcgggct ctcttaagct tttccacag cattgggaaa ctttactgac 1500

```

RAW SEQUENCE LISTING

DATE: 09/23/2002

PATENT APPLICATION: US/10/005,337A

TIME: 10:56:26

Input Set : A:\Substitute Sequence Listing.txt

Output Set: N:\CRF3\09232002\J005337A.raw

```

55 agcatccaag ttgtgcttct gctaagaact ggactcacat ctctctgtgc atcacttcgg 1560
56 cccgttttgg ggtagatcct ctgattagcc ttcagattta gaacacggtg agcctgtggt 1620
57 gcaactaatta tggccagtga caccatagag tcaaagtga ttactgaatg ctttcaattt 1680
58 ctocctaatgc tggtaacgat gcatgtcaca gggccatttt agctgcagac atcactccag 1740
59 agaattccaa acagatagag acaagtggca cccagaccca tctccttccc ctggggctga 1800
60 ttatccccag aaataggatg tcccaaagca acacttccca gccaaactgga gtgctgataa 1860
61 gtccagttat cagaaagata tggctgtaag tgtgatgcac agtgcttgca ttttcttgat 1920
62 acgttagtca tatgagagct gacaaagaag gaaaaagagc agcgatgtgg tgcaatatta 1980
63 acaggcagct gtcccctggc ttcccgatac gtgggatgac tcgcattgct gagcgggtgtg 2040
64 gtcactgcca aaggaatgac cctctcacat ttcttctga ttgcatacgc ccgcggccag 2100
65 cttgtcatct cctcttggg cttcccagac actaagtctg gaatgaaaat tcacctgcct 2160
66 ctgaattggc cactggtggg ggcaggggtg tgacttggct tcccaggctg gaagattatc 2220
67 tcaccacagc ctagctatat aacgggctgg tgtggagggg ctccacaggg ccagttccag 2280
68 gggttcatcc acaagagaga aaaacataga ctogaggctc agggagcttg catgcctgca 2340
69 ggtcggaggc caccatgg                                     2358
72 <210> SEQ ID NO: 2
73 <211> LENGTH: 2074
74 <212> TYPE: DNA
75 <213> ORGANISM: Homo sapiens
77 <400> SEQUENCE: 2
78 ctgcagcaag ttacttaatg ttttttgct cagcatcctc tctgtaaaat gagagcatta 60
79 gtcttgcctc aacttcgagg gcatggacag ctctgggatt tcatatccaa gacccttaaa 120
80 catcccacag tcttcccccc aaacacttct cctcctaata cctccctcag tttgggtcag 180
81 gcctggaaca aaaaggcata cgaaatggtg gaaaaagtgt ccactgactac ttctgactaa 240
82 gatgaagaga ccaatgaaaa tagtaatgac tctgtttgct tcagcaggac atatactaaa 300
83 ataggagcta tacaagaag attagcatgg actctgtgca agaatgacac acaaatttgt 360
84 gaaacattcc atatattaaa aataaataaa taataaagag aaaaggaaaa aattaaaaag 420
85 aaaatagtga tagctgtgtc catctcaaag aaaagcccag gagatttctt ttatttaccc 480
86 cctttaagat agaataatag gagaccgga catatgatac aggaggtaact gggagggtcc 540
87 ctctttgtca atgttttgtc ttgggggtgg gagtcgatgt cttctcaaag tttcagaaac 600
88 accatccact gactgagcat tcaaggggca agaggagaat ggcagccaca tttgttgatt 660
89 gggtagagtt ggggagaaat agacacacaa aggtcaaaca taacttcta attaacactt 720
90 cctccattc acaattccct tctccattc ttctctctg tcttttacts akaraaacc 780
91 agtttttctt gaaactataa aaatacccc agtatgttta cataatttac acctcaaaga 840
92 ttagaaacca gaaatagaga ccttttcaac ccttccggaa gcaaagtga ttatccctcc 900
93 agccacgtgt ctcaaattct gatgcatcag aatcatctgg gtgctttkaa attcaagatg 960
94 attcctacga gttaccataa atcaactcag aattccctgg agtggggcca gggatctgta 1020
95 tttctgacaa gctcccacag gtgattcctt tcccacagc atttgagaac ttcagctcaa 1080
96 tgacctaatc agagtctgct cattgcta atctgggtctc atttttbtca tatatatata 1140
97 tagtatttgt ggtagagatg ggattttgcc atgttgccca ggctagtatt gaactcctaa 1200
98 gctaagcaat ctctctgtct ctgctccca aatgtttggg attacaggtg taagccactg 1260
99 caccggctg atagctggtt tcatttactc tatttcttga ccactctgat ccattttgaa 1320
100 gtaaaaaatgc tccaattatt atgtgtttt agaacaagg aagcatgtca tgtgctaattg 1380
101 gccagtga caataaaaga aaagtgcatt actgaatgct ttaaatgtct tataatgatg 1440
102 gtaagggtgg atgtcatggg gctattttag cccagacatc actccaaaga attccaaaca 1500
103 gatatagaca agtgcttita gggccagat ccttccccct caggctgttt acccagggaa 1560
104 taggatgtcc tgggacaagt ttcccctaag tgaagtgttg ataagtctgc ttatcagaaa 1620
105 gatattactg ggggtgtgat atgtagggca tctacatttt cttgatagggt agtcatatga 1680
106 aagctgacaa agaaaaaaag ggcagtgatg tgggtgcaatg tcaacagaca gctgtcccct 1740

```

RAW SEQUENCE LISTING

DATE: 09/23/2002

PATENT APPLICATION: US/10/005,337A

TIME: 10:56:26

Input Set : A:\Substitute Sequence Listing.txt

Output Set: N:\CRF3\09232002\J005337A.raw

```

107 gactcttgac aaataggatg acttgcatg ctgagcgatg tgatcaccac caaaggaatg 1800
108 gccctctcac atttcttcct gattcacata ttcagcaggg ttagcttgtc ctcccctccc 1860
109 tcttcagctt cccagacact gagtctggaa tgaaaattca cctgcctctg agttggctcc 1920
110 taatgggggc gggagtgtta ctgcggttcc caggttgga gattatctca cccggcccca 1980
111 gctatataag ctgaccggtg tggagggggc cagcagggcc aactccaggg attccttcca 2040
112 cgacagaaaa acatacaaga ctcttcagc caac 2074
115 <210> SEQ ID NO: 3
116 <211> LENGTH: 750
117 <212> TYPE: PRT
118 <213> ORGANISM: Homo sapiens
120 <400> SEQUENCE: 3
121 Met Gly Glu Thr Leu Gly Asp Ser Pro Ile Asp Pro Glu Ser Asp Ser
122 1 5 10 15
124 Phe Thr Asp Thr Leu Ser Ala Asn Ile Ser Gln Glu Met Thr Met Val
125 20 25 30
127 Asp Thr Glu Met Pro Phe Trp Pro Thr Asn Phe Gly Ile Ser Ser Val
128 35 40 45
130 Asp Leu Ser Val Met Glu Asp His Ser His Ser Phe Asp Ile Lys Pro
131 50 55 60
133 Phe Thr Thr Val Asp Phe Ser Ser Ile Ser Thr Pro His Tyr Glu Asp
134 65 70 75 80
136 Ile Pro Phe Thr Arg Thr Asp Pro Val Val Ala Asp Tyr Lys Tyr Asp
137 85 90 95
139 Leu Lys Leu Gln Glu Tyr Gln Ser Ala Ile Lys Val Glu Pro Ala Ser
140 100 105 110
142 Pro Pro Tyr Tyr Ser Glu Lys Thr Gln Leu Tyr Asn Lys Pro His Glu
143 115 120 125
145 Glu Pro Ser Asn Ser Leu Met Ala Ile Glu Cys Arg Val Cys Gly Asp
146 130 135 140
148 Lys Ala Ser Gly Phe His Tyr Gly Val His Ala Cys Glu Gly Cys Lys
149 145 150 155 160
151 Gly Phe Phe Arg Arg Thr Ile Arg Leu Lys Leu Ile Tyr Asp Arg Cys
152 165 170 175
154 Asp Leu Asn Cys Arg Ile His Lys Lys Ser Arg Asn Lys Cys Gln Tyr
155 180 185 190
157 Cys Arg Phe Gln Lys Cys Leu Ala Val Gly Met Ser His Asn Ala Ile
158 195 200 205
160 Arg Phe Gly Arg Met Pro Gln Ala Glu Lys Glu Lys Leu Leu Ala Glu
161 210 215 220
163 Ile Ser Ser Asp Ile Asp Gln Leu Asn Pro Glu Ser Ala Asp Leu Arg
164 225 230 235 240
166 Ala Leu Ala Lys His Leu Tyr Asp Ser Tyr Ile Lys Ser Phe Pro Leu
167 245 250 255
169 Thr Lys Ala Lys Ala Arg Ala Ile Leu Thr Gly Lys Thr Thr Asp Lys
170 260 265 270
172 Ser Pro Phe Val Ile Tyr Asp Met Asn Ser Leu Met Met Gly Glu Asp
173 275 280 285
175 Lys Ile Lys Phe Lys His Ile Thr Pro Leu Gln Glu Gln Ser Lys Glu
176 290 295 300

```

RAW SEQUENCE LISTING

DATE: 09/23/2002

PATENT APPLICATION: US/10/005,337A

TIME: 10:56:26

Input Set : A:\Substitute Sequence Listing.txt

Output Set: N:\CRF3\09232002\J005337A.raw

```

178 Val Ala Ile Arg Ile Phe Gln Gly Cys Gln Phe Arg Ser Val Glu Ala
179 305 310 315 320
181 Val Gln Glu Ile Thr Glu Tyr Ala Lys Ser Ile Pro Gly Phe Val Asn
182 325 330 335
184 Leu Asp Leu Asn Asp Gln Val Thr Leu Lys Tyr Gly Val His Glu
185 340 345 350
187 Ile Ile Tyr Thr Met Leu Ala Ser Leu Met Asn Lys Asp Gly Val Leu
188 355 360 365
190 Ile Ser Glu Gly Gln Gly Phe Met Thr Arg Glu Phe Leu Lys Ser Leu
191 370 375 380
193 Arg Lys Pro Phe Gly Asp Phe Met Glu Pro Lys Phe Glu Phe Ala Val
194 385 390 395 400
196 Lys Phe Asn Ala Leu Glu Leu Asp Asp Ser Asp Leu Ala Ile Phe Ile
197 405 410 415
199 Ala Val Ile Ile Leu Ser Gly Asp Arg Pro Gly Leu Leu Asn Val Lys
200 420 425 430
202 Pro Ile Glu Asp Ile Gln Asp Asn Leu Leu Gln Ala Leu Glu Leu Gln
203 435 440 445
205 Leu Lys Leu Asn His Pro Glu Ser Ser Gln Leu Phe Ala Lys Leu Leu
206 450 455 460
208 Gln Lys Met Thr Asp Leu Arg Gln Ile Val Thr Glu His Val Gln Leu
209 465 470 475 480
211 Leu Gln Val Ile Lys Lys Thr Glu Thr Asp Met Ser Leu His Pro Leu
212 485 490 495
214 Leu Gln Glu Ile Tyr Lys Asp Leu Tyr Ala Trp Ala Ile Leu Thr Gly
215 500 505 510
217 Lys Thr Thr Asp Lys Ser Pro Phe Val Ile Tyr Asp Met Asn Ser Leu
218 515 520 525
220 Met Met Gly Glu Asp Lys Ile Lys Phe Lys His Ile Thr Pro Leu Gln
221 530 535 540
223 Glu Gln Ser Lys Glu Val Ala Ile Arg Ile Phe Gln Gly Cys Gln Phe
224 545 550 555 560
226 Arg Ser Val Glu Ala Val Gln Glu Ile Thr Glu Tyr Ala Lys Ser Ile
227 565 570 575
229 Pro Gly Phe Val Asn Leu Asp Leu Asn Asp Gln Val Thr Leu Leu Lys
230 580 585 590
232 Tyr Gly Val His Glu Ile Ile Tyr Thr Met Leu Ala Ser Leu Met Asn
233 595 600 605
235 Lys Asp Gly Val Leu Ile Ser Glu Gly Gln Gly Phe Met Thr Arg Glu
236 610 615 620
238 Phe Leu Lys Ser Leu Arg Lys Pro Phe Gly Asp Phe Met Glu Pro Lys
239 625 630 635 640
241 Phe Glu Phe Ala Val Lys Phe Asn Ala Leu Glu Leu Asp Asp Ser Asp
242 645 650 655
244 Leu Ala Ile Phe Ile Ala Val Ile Ile Leu Ser Gly Asp Arg Pro Gly
245 660 665 670
247 Leu Leu Asn Val Lys Pro Ile Glu Asp Ile Gln Asp Asn Leu Leu Gln
248 675 680 685
250 Ala Leu Glu Leu Gln Leu Lys Leu Asn His Pro Glu Ser Ser Gln Leu

```

RAW SEQUENCE LISTING

DATE: 09/23/2002

PATENT APPLICATION: US/10/005,337A

TIME: 10:56:26

Input Set : A:\Substitute Sequence Listing.txt

Output Set: N:\CRF3\09232002\J005337A.raw

```

251      690      695      700
253 Phe Ala Lys Leu Leu Gln Lys Met Thr Asp Leu Arg Gln Ile Val Thr
254 705      710      715      720
256 Glu His Val Gln Leu Leu Gln Val Ile Lys Lys Thr Glu Thr Asp Met
257      725      730      735
259 Ser Leu His Pro Leu Leu Gln Glu Ile Tyr Lys Asp Leu Tyr
260      740      745      750
263 <210> SEQ ID NO: 4
264 <211> LENGTH: 30
265 <212> TYPE: DNA
266 <213> ORGANISM: Artificial sequence
268 <220> FEATURE:
269 <223> OTHER INFORMATION: PCR Primer
271 <400> SEQUENCE: 4
272 ggcgatttaa ataatgtagt cttatgcaat 30
275 <210> SEQ ID NO: 5
276 <211> LENGTH: 31
277 <212> TYPE: DNA
278 <213> ORGANISM: Artificial sequence
280 <220> FEATURE:
281 <223> OTHER INFORMATION: PCR Primer
283 <400> SEQUENCE: 5
284 ggggtctaga aggtgcacac caatgtggtg a 31

```

VERIFICATION SUMMARY

DATE: 09/23/2002

PATENT APPLICATION: US/10/005,337A

TIME: 10:56:27

Input Set : A:\Substitute Sequence Listing.txt

Output Set: N:\CRF3\09232002\J005337A.raw